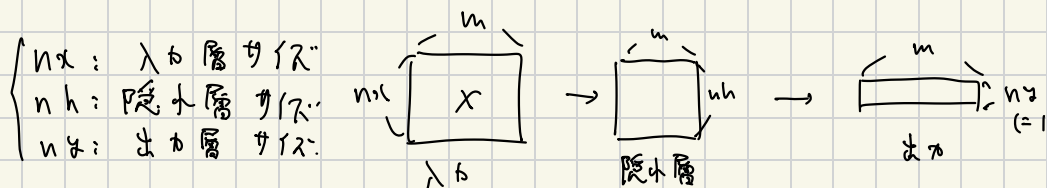


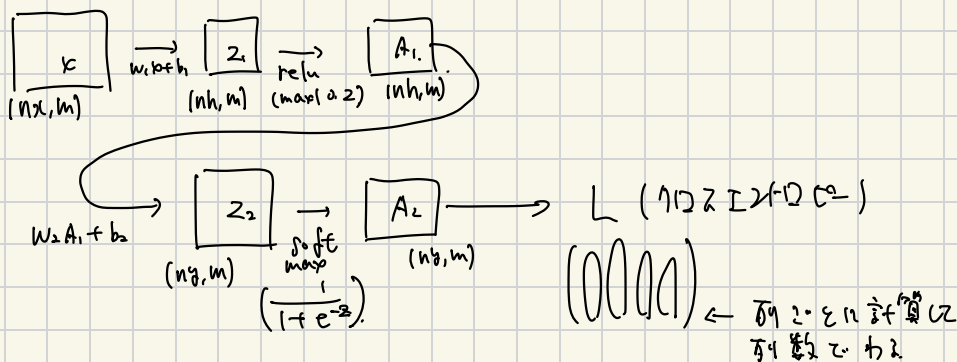
二層の場合



$$\begin{cases} W_1 = \text{Mat}(n_x, n_h) \\ W_2 = \text{Mat}(n_h, n_y) \\ b_1 = n_h \text{ 次元ベクトル} \\ b_2 = n_y \text{ 次元ベクトル} \end{cases} \quad \text{バイアスはベクトルなので}$$

各層に1回ずつ、 m 回定ておす。

順伝播



逆伝播

$$y \in Y, L = -\{y \log(A_2) + (1-y) \log(1-A_2)\}$$

$$\begin{aligned} \frac{dL}{dA_2} &= -\left\{ y \frac{1}{A_2} + (1-y) \frac{-1}{1-A_2} \right\} \\ &= -\frac{y(1-A_2) - A_2(1-y)}{A_2(1-A_2)} = \frac{A_2 - y}{A_2(1-A_2)} \end{aligned}$$

$$A_2 = \frac{1}{1 + e^{-z_2}}$$

$$\frac{dA_2}{dz_2} = \frac{-(-e^{-z_2})}{(1+e^{-z_2})^2} = A_2(1-A_2)$$

$$\frac{dL}{dz_2} = \frac{dL}{dA_2} \cdot \frac{dA_2}{dz_2} = \frac{A_2 - y}{A_2(1-A_2)} \cdot A_2(1-A_2) = A_2 - y$$

$$\rightarrow \frac{dL}{dz_2} = A_2 - y \quad (= dz_2)$$

$$\frac{dz_2}{dw_2} = A_1 \quad \frac{dL}{dw_2} = \frac{dL}{dz_2} \cdot \frac{dz_2}{dw_2} = dz_2 \cdot A_1 \quad \leftarrow \begin{array}{l} \text{次元合わせのため} \\ \text{に転置をとっている} \end{array}$$

$$\rightarrow dW_2 = \frac{1}{m} (dZ_2 \cdot A_2^T)$$

(← 平均値とL2計算を平均化する)

$$\frac{dL}{db_2} = \frac{dL}{dZ_2} \cdot \frac{dZ_2}{db_2} = \frac{dZ_2}{(n_y, m) \text{ 次元}} \quad \leftarrow \text{横方向に和をとる (m回足してn分割する)}$$

ここで $\frac{dL}{dw_1}, \frac{dL}{db_2}$ (パラメータごとの損失関数のセアン)

同様にして $\frac{dL}{dw_1}, \frac{dL}{db_2}$ を求める。

これを求めるのが目標。

1パラメータ更新

$$\begin{cases} w_1 = w_1 - lr \times \frac{dL}{dw_1} \\ w_2 = w_2 - lr \times \frac{dL}{dw_2} \\ b_1 = b_1 - lr \times \frac{dL}{db_1} \\ b_2 = b_2 - lr \times \frac{dL}{db_2} \end{cases}$$